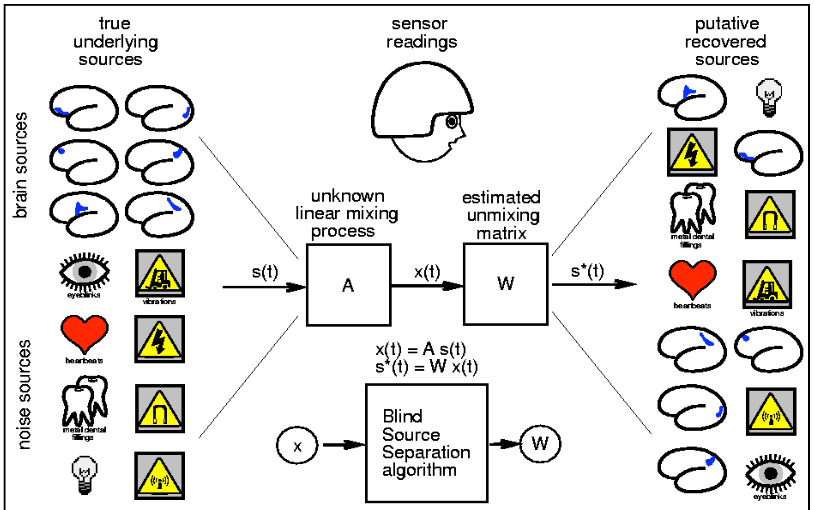
Carl Leichter (P-21); Mentor: Robert Kraus, Jr. (P-21):

**Issues in Using Independent Component Analysis (ICA) and Principal Component Analysis (PCA) for Noise Reduction in the Analysis of Single-Trial Superconducting Image Surface (SIS) Magnetoencephalographic (MEG) Data**

The goals of the MEG program is to develop, test, and evaluate novel superconducting quantum interference device (SQUID) sensor concepts and devices, new models of electromagnetic sources in the brain, numerical techniques, and computational models for functional imaging of the human brain using MEG with other methods. MEG directly measures a physical effect of neuronal currents with a temporal resolution not limited by the sluggish vascular response, unlike positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) that measure hemodynamic changes (changes in blood flow) associated with brain activity.

High temporal resolution is particularly important for studying neurological disorders such as epilepsy where temporal information is a major diagnostic. It is also important for fundamental studies of synchronization and oscillatory brain activity. The whole-head MEG system is based on the Los Alamos-patented principle of superconducting image surface (SIS) gradiometry where magnetic sources are imaged on the surface and magnetometers near this surface sense the combined fields as if the sensors were gradiometers.




The icons at left represent the true sources of signal (both brain activity and environmental interference). The MEG machine sees them all mixed together (box A). Our task is to computationally separate these signals (box W) to isolate the signals from the brain that we wish to study, represented by the icons on the right. This task is greatly complicated by the fact that we don't actually see the signals indicated by the left-hand icons. We only see the mixed initial condition—and we cannot be certain exactly how many signals there are nor what their precise individual characteristics are.

A significant problem for our system is that we live in a world awash in electromagnetic chatter. Modern buildings are saturated with these spurious signals. The SQUID detectors in the MEG system are so sensitive that this noise threatens to overwhelm them. The current experimental MEG apparatus resides within a room that is electromagnetically shielded from the outside world. That system can better focus on the signals coming from the person under study. However, the human body also generates extraneous electrical signals that must also be separated out and computationally eliminated.

Our goal with this project is even harder: to analyze the MEG data in a manner that will allow us to determine what is happening in the brain *while* it is happening—that is, online, real-time analysis. We cannot isolate our test subject in a shielded room if we want to accurately measure brain activity in a real-world situation. The basic issue is one of signal versus noise. We need to identify, isolate, and then computationally eliminate all of the external noise of the environment, as well as electromagnetic noise produced by the human body, to study that portion of the brain's electrical activity indicative of the subject's thought processes.

The electrical activity that we measure with our sensor is a *mixture* of the brain signals that we want with the noise that we don't want. We need to find a way to separate the signal from the noise. PCA and ICA are two methods of multichannel (or multidimensional) signal decomposition. We have tried to use these signal-analysis techniques to decompose these signal and noise mixtures to separate the signal from the noise. In their simplest form, PCA and ICA can both be understood from the point of view of *projection pursuit*. Imagine that our multidimensional data are distributed like a cloud in our data space. A one-dimensional vector (a straight line) is placed next to our data cloud and we look at the "shadow" that the data cloud casts onto this vector. This shadow is the *projection* of the data onto the vector. We move the vector around and examine the characteristics of the projection as the vector is moved. In the case of PCA, we are interested in the variance of this projection. We are seeking (or *pursuing*) the projection that has the maximum variance. Once we have achieved this optimization, we *extract* this projection as the first principle component (the component with maximal variance). By *deflation,* we remove the first principle component from the data. We can then repeat this pursuit, extraction, and deflation process with the leftover (residual) data to find the second principle component, the third principle component, and so on until we have extracted as many components as there are channels of data. ICA differs from PCA in that instead of pursuing vectors that have maximum variance, we are pursuing vectors that have minimum gaussianity.

Why do we seek minimum gaussianity with ICA? There is a principle in statistics called the Central Limit Theorem (sometimes referred to as the "law of large numbers"). This is best illustrated by flipping coins. If you flip one coin, there is a 50/50 chance that it will land heads (or tails). It can only be one or the other, and the *probability* of a given outcome for a single trial is always 50% or 0.5. This is called a *binomial distribution*. Now, suppose we were to flip 1,000 coins at once, what are the possible outcomes? We could get all heads (or all tails) or some combination of heads and tails. As you can imagine, it is very unusual for every coin to land as heads (or tails). We describe such an outcome as having a low probability. The outcome with the highest probability is that half of the coins will land as heads and the other half will land as tails. If you were to look at the distribution (graph) of the number of heads/tails versus the probability of that number of heads/tails, it would look a "bell" curve. The highest (most probable) part of the bell curve is the most likely outcome and that is the half heads/half tails outcome. The lowest parts of the bell curve (the two ends or "tails") represent the least likely outcome: all tails on one end and all heads on the other end.

Another name for the bell curve is a *gaussian* or normal distribution. If you'll recall, the individual coin toss had a binomial distribution. But, when we *mixed* a large number of coin tosses together, the resulting distribution tended towards a gaussian distribution. This is the law of large numbers: no matter what the original distribution was, when you start to mix things together, the mixture begins to look like a gaussian distribution. The Central Limit Theorem is a mathematical expression of this law. Its application to signal mixtures is analogous: no matter what the original statistical characteristics of your original signals were, a mixture of those signals will have gaussian statistics. So, if we want to *unmix* that signal mixture, we are looking for projections that are as divergent from the gaussian distribution as possible. There is one thing you have to watch out for—this technique only works if you have AT MOST one component that has a gaussian distribution. Obviously, if you have a mixture of things that are gaussian to start with, you will not be able to find projections that diverge from gaussianity.

During our study, we found that attempting to remove noise using the PCA method can cause signal distortion: both a loss of signal and the insertion of a spurious signal. Removing noise with the ICA method requires order estimation. If we estimate that there are too few components or too many components, we will fail to extract statistically independent components, and the signal is mixed with noise. In principle, PCA *cannot* guarantee this separation. But, assuming that the noise is statistically independent from the signal, ICA is designed to give *precisely* this kind of separation. ICA will solve all our separation problems…if we can accomplish accurate order separations.